

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Patent Application of	)	
	)	
Yoko FUJIWARA	)	Group Art Unit: Unassigned
	)	
Application No.: Unassigned	)	Examiner: Unassigned
	)	
Filed: June 7, 2001	)	
	)	
For: IMAGE PROCESSOR FOR	)	
CHARACTER RECOGNITION	)	
	)	
	)	
	)	



CLAIM FOR CONVENTION PRIORITY

Assistant Commissioner for Patents  
Washington, D.C. 20231

Sir:

The benefit of the filing date of the following prior foreign application in the following foreign country is hereby requested, and the right of priority provided in 35 U.S.C. § 119 is hereby claimed:

Japanese Patent Application No. 2000-173727

Filed: June 9, 2000

In support of this claim, enclosed is a certified copy of said prior foreign application. Said prior foreign application was referred to in the oath or declaration. Acknowledgment of receipt of the certified copy is requested.

Respectfully submitted,

BURNS, DOANE, SWECKER & MATHIS, L.L.P.

Date: June 7, 2001

By:

Platon N. Mandros  
Registration No. 22,124

P.O. Box 1404  
Alexandria, Virginia 22313-1404  
(703) 836-6620

日本国特許庁  
PATENT OFFICE  
JAPANESE GOVERNMENT

10/07/90  
09/875081  
10903 U.S. PTO

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出願年月日  
Date of Application:

2000年 6月 9日

出願番号  
Application Number:

特願2000-173727

願 人  
Applicant(s):

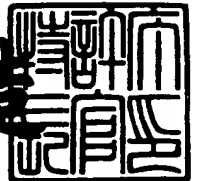
ミノルタ株式会社

CERTIFIED COPY OF  
PRIORITY DOCUMENT

2001年 3月 9日

特許庁長官  
Commissioner,  
Patent Office

及川耕造



出証番号 出証特2001-3016984

【書類名】 特許願

【整理番号】 169534

【提出日】 平成12年 6月 9日

【あて先】 特許庁長官殿

【国際特許分類】 G06T 5/00

【発明者】

【住所又は居所】 大阪府大阪市中央区安土町二丁目3番13号大阪国際ビル ミノルタ株式会社内

【氏名】 藤原 葉子

【特許出願人】

【識別番号】 000006079

【住所又は居所】 大阪府大阪市中央区安土町二丁目3番13号大阪国際ビル

【氏名又は名称】 ミノルタ株式会社

【代理人】

【識別番号】 100062144

【弁理士】

【氏名又は名称】 青山 葆

【選任した代理人】

【識別番号】 100086405

【弁理士】

【氏名又は名称】 河宮 治

【選任した代理人】

【識別番号】 100098280

【弁理士】

【氏名又は名称】 石野 正弘

【手数料の表示】

【予納台帳番号】 013262

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【包括委任状番号】 9808001

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 画像処理のための装置、方法及び記録媒体

【特許請求の範囲】

【請求項 1】 画像データ中の文字画像データを文字コードデータに変換する画像処理装置において、

背景画像上に文字画像が存在する画像データから文字画像データを抽出する文字画像データ抽出部と、

抽出された文字画像データを文字コードデータに変換する変換部と、

背景画像上の文字画像を除去する文字画像除去部と、

文字コードデータと文字画像除去後の画像データとを合成する合成部と

を有することを特徴とする画像処理装置。

【請求項 2】 請求項 1 に記載された画像処理装置において、

文字画像除去部は、背景画像上の文字画像部分の画像データを周辺背景画像データに応じて補正することを特徴とする画像処理装置。

【請求項 3】 請求項 1 に記載された画像処理装置において、

前記の変換部は、色変化のある文字画像データを文字コードデータに変換しないことを特徴とする画像処理装置。

【請求項 4】 請求項 1 に記載された画像処理装置において、

前記の文字データ抽出部は、1 文字ずつ文字コードデータを抽出することを特徴とする画像処理装置。

【請求項 5】 請求項 1 に記載された画像処理装置において、

前記の文字データ抽出部は、単語単位で文字コードデータを抽出することを特徴とする画像処理装置。

【請求項 6】 画像データ中の文字画像データを文字コードデータに変換する画像処理方法であって、

背景画像上に文字画像が存在する画像データから文字画像データを抽出するステップと、

抽出された文字画像データを文字コードデータに変換するステップと、

背景画像上の文字画像部分の画像データを周辺背景画像データに応じて補正す

るステップと、

文字コードデータと補正後の画像データとを合成するステップと  
からなる画像処理方法。

【請求項 7】 背景画像上に文字画像が存在する画像データから文字画像データを抽出するステップと、

抽出された文字画像データを文字コードデータに変換するステップと、

背景画像上の文字画像部分の画像データを周辺背景画像データに応じて補正するステップと、

文字コードデータと補正後の画像データとを合成するステップと

からなる、画像データ中の文字画像データを文字コードデータに変換する画像処理プログラムを記録するコンピュータ読み取り可能な記録媒体。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、画像データ中の文字画像データを文字コードデータに変換する文字認識に関する。

【0002】

【従来の技術】

従来から、文字認識(OCR)装置による文字認識技術では、テキスト画像の文字を認識するだけではなく、原稿のレイアウトも認識し、文字以外の画像データはビットマップ形式の画像データとしてレイアウトに応じた所定位置に配置し、1つのファイルとして出力するものが存在している。

【0003】

【発明が解決しようとする課題】

しかし、背景画像上に文字が描かれている原稿において、その文字画像データを文字コードデータに変換した後で、元の画像データと合成することは困難であった。なぜなら、元の画像データで使用されている文字のフォントや配置と文字コードデータに変換した後の文字のフォントや配置とが必ずしも一致しないからである。

従来の文字認識技術においては、例えば、図 1 に示すように、原稿上の文字画像データを文字コードデータに変換するだけで、背景画像データは出力しないものや、図 2 に示すように、原稿の画像データと変換後の文字コードデータとを重ねて出力するものが存在していた。しかし、前者では、背景画像がなくなっており、情報が欠落してしまうし、後者では、文字画像データと文字コードデータによる出力画像がずれており、出力画像が見づらいものになってしまう。

#### 【0004】

本発明の目的は、背景のある文字をうまく再現する画像処理のための装置、方法および記録媒体を提供することである。

#### 【0005】

##### 【課題を解決するための手段】

本発明に係る画像処理装置は、画像データ中の文字画像データを文字コードデータに変換する画像処理装置であって、背景画像上に文字画像が存在する画像データから、文字画像データを抽出する文字画像データ抽出部と、抽出された文字画像データを文字コードデータに変換する変換部と、背景画像上の文字画像を除去する文字画像除去部と、文字コードデータと文字画像除去後の画像データとを合成する合成部とを有する。

前記の画像処理装置において、好ましくは、文字画像除去部は、背景画像上の文字画像部分の画像データを周辺背景画像データに応じて補正する。

前記の画像処理装置において、好ましくは、前記の変換部は、色変化のある文字画像データを文字コードデータに変換しない。

前記の画像処理装置において、前記の文字データ抽出部は、たとえば、1文字づつ文字コードデータを抽出する。

前記の画像処理装置において、前記の文字データ抽出部は、たとえば、単語単位で文字コードデータを抽出する。

#### 【0006】

本発明に係る画像処理方法は、画像データ中の文字画像データを文字コードデータに変換する画像処理方法であって、背景画像上に文字画像が存在する画像データから、文字画像データを抽出するステップと、抽出された文字画像データを

文字コードデータに変換するステップと、背景画像上の文字画像部分の画像データを周辺背景画像データに応じて補正するステップと、文字コードデータと補正後の画像データとを合成するステップとからなる。

## 【 0 0 0 7 】

本発明に係るコンピュータ読み取りが可能な記録媒体は、背景画像上に文字画像が存在する画像データから、文字画像データを抽出するステップと、抽出された文字画像データを文字コードデータに変換するステップと、背景画像上の文字画像部分の画像データを周辺背景画像データに応じて補正するステップと、文字コードデータと補正後の画像データとを合成するステップとからなる、画像データ中の文字画像データを文字コードデータに変換する画像処理プログラムを記録する。

## 【 0 0 0 8 】

## 【発明の実施の形態】

以下、添付の図面を参照して発明の実施の形態を説明する。なお、図面において、同じ参照記号は同一または同等のものを示す。

図 3 は、発明の 1 実施形態におけるシステム構成を示す。ネットワーク 1 0 0 に、文字認識機能を有する文字認識処理装置としてコンピュータ 2 0 0、画像出力装置としてプリンタ 3 0 0、画像入力装置としてスキャナ 4 0 0 などが接続されている。本実施形態では、コンピュータ 2 0 0 が文字認識機能を有しているが、文字認識機能を画像出力装置または画像入力装置に持たせてもよい。また、画像出力装置は、プリンタ 3 0 0 に限らず、ディスプレイであってもよい。さらに、画像入力装置は、スキャナ 4 0 0 に限らず、デジタルカメラ、フィルムスキャナなどのデジタル画像を入力する機器であってもよい。

## 【 0 0 0 9 】

コンピュータ 2 0 0 には、文書認識処理部 2 1 0 と、スキャナドライバ 2 3 0、プリンタドライバ 2 4 0 などの周辺機器を制御するソフトが搭載されている。例えば、スキャナドライバ 2 3 0 を用いて、コンピュータ 2 0 0 上の画面から、スキャナ 4 0 0 の読取条件の設定や読み取り開始が指示できる。

## 【 0 0 1 0 】



なお、本実施形態ではスキャナ 4 0 0 で読み取って得られた画像データに対してコンピュータ 2 0 0 で処理を行っている。しかし、同様の処理をスキャナ 4 0 0 内で行ってもよいし、プリンタ 3 0 0 で行うようにしてもよい。また、スタンダロンのデジタル複写機のような形態にも適用できる。

#### 【 0 0 1 1 】

文書認識処理部 2 1 0 では、背景画像上に文字画像が存在する画像データから文字画像データを抽出し、抽出された文字画像データを文字コードデータに変換する。また、背景画像上の文字画像部分の画像データを周辺背景画像データに応じて補正する。そして、文字コードデータと補正後の画像データとを合成する。具体的には、文字認識部 2 1 2 は、画像を読み取って得られた画像データから文字画像データを認識し、文字画像データを文字コードに変換して出力する。文字除去部 2 1 4 は、認識した文字コードデータを元の画像データから除去し、画像認識部 2 1 6 は、文字画像データが除去された画像データから、さらに余白(白地で何も描かれていない領域)以外の写真・図柄などの画像データを認識する。画像合成部 2 1 8 は、文字コードデータと、画像認識部 2 1 6 で認識された画像データとを合成する。これにより、元の原稿に含まれる文字画像データと変換された文字コードデータが重なることがない。また、原稿に含まれる背景を残したまま、原稿に含まれる文字画像データにおいて、文字コードデータを用いて、文章の編集・加工が容易であり、また、検索キーワードとしても使用可能になる。

#### 【 0 0 1 2 】

図 4 は、文書認識処理部 2 1 0 の処理を模式的に示したものである。左側の図は、背景画像の上に文字が描かれている原稿を示し、この原稿に対して、右下の図は、原稿上の文字画像データを文字認識部 2 1 2 で文字コードデータに変換した画像データを示し、右上の図は、ビットマップ画像認識部 2 1 6 で認識した画像から原稿上の文字画像データを文字除去部 2 1 4 で除去したビットマップ画像データを示す。

#### 【 0 0 1 3 】

図 5 は、文書認識処理部 2 1 0 における処理のフローチャートである。まずスキャナ 4 0 0 により、原稿を読み取って得られた RGB 成分で構成された画像デ

ータを、個々のスキャナデバイスに依存しない L a b 成分の画像データへと色空間変換をする (S 1 0)。

#### 【 0 0 1 4 】

次に、色空間が変換された画像データに対して、文字認識部 2 1 2 での認識率を上げるため、OCR 前処理を行う (S 1 2)。具体的には、L a b 成分の画像データに対し、文字画像データが黒、背景画像データが白となるように、2 値化処理を行う。ここで、L a b 成分の画像データは別の場所に保存しておく。さらに OCR 前処理では、画像データに含まれる孤立点などのノイズ除去、原稿読取時の原稿の傾きによる画像データの傾きの補正、つぶれ文字やかすれ文字の補正なども行う。ここでは、L a b 成分の画像データに対して OCR 前処理を行っているが、色空間変換前の R G B 成分の画像データに対して OCR 前処理を行ってもよい。

#### 【 0 0 1 5 】

次に、文字認識部 2 1 2 で、画像データに含まれる文字画像データで構成される行、及び、各文字画像データを画像データから切り出し、各文字画像データについて文字コード、フォント、フォントサイズなどが認識される (S 1 4)。文字コードデータの抽出は、たとえば 1 文字ずつ行うが、単語単位で行ってもよい。フォント及びフォントサイズについては、文書認識処理部 2 1 0 が有するフォントデータ群から最も類似度の高いフォント及びフォントサイズが選択される。また、文字認識部 2 1 2 は、認識された文字画像データの位置データとして、図 6 に示すような、原点 (0, 0) を画像データの左上端とする文字画像データを含む矩形の位置座標 (X 1, Y 1)、(X 2, Y 2) を出力する。

#### 【 0 0 1 6 】

なお、1 つの文字内で色が変化した場合は、その文字は文字以外のものを誤認識したものとみなし、コード化は行わない。すなわち、色変化のある文字画像データは文字コードデータに変換しない。上述の OCR 前処理 (S 1 2) において L a b 成分の画像データに対して 2 値化処理を行っているが、L a b 成分の画像データは別の場所に保存されており、その保存されている画像データから文字の色を検出する。具体的には、まず、図 7 に示す 3 × 3 最大値フィルタを用いて

、文字認識のための入力モノクロ2値画像について近傍3×3画素の最大値をとるフィルタ処理を行う。これにより、文字部分を収縮させる。これは、文字のエッジ部の影響を除去するためである。こうして2値画像を収縮処理した画像と、文字認識により得られた文字外接矩形とのANDをとり、その画像の画素についての平均データを求めてその文字の色とする。即ち、j番目の文字のL a b空間で表したカラー画像データを $C_j$ とすると、

【数1】

$$C_j = \left( \frac{\sum_{i=1}^N L_i}{N}, \frac{\sum_{i=1}^N a_i}{N}, \frac{\sum_{i=1}^N b_i}{N} \right)$$

ここに、Nはj番目の文字の画素数である。

【0 0 1 7】

色が変化したか否かの判定は、収縮処理した文字の外接矩形を図8に破線で示すように4つに分割し、それぞれの領域でのL a b値の平均を求めて比較することにより行う。4つの領域の平均L a b値が一定値以上異なる場合は、色が変わったと判別される。4つの領域でのL a b値を $C_{j1} \sim C_{j4}$ とすると、

$$C_{j1} = (L_{j1}, a_{j1}, b_{j1})$$

$$C_{j2} = (L_{j2}, a_{j2}, b_{j2})$$

$$C_{j3} = (L_{j3}, a_{j3}, b_{j3})$$

$$C_{j4} = (L_{j4}, a_{j4}, b_{j4})$$

で表される。次の条件

$$|L_{jn} - L_{jm}| \geq k_1 \quad \text{または}$$

$$|a_{jn} - a_{jm}| \geq k_2 \quad \text{または}$$

$$|b_{jn} - b_{jm}| \geq k_3$$

(ここに、 $n = 1, 2, 3, 4$ 、 $m = 1, 2, 3, 4$ であり、 $k_1$ 、 $k_2$ 、 $k_3$ は定数である)を満たす場合、色が変わったと判別される。前述の1文字の平均画像データ $C_j$ は、実際は、これら4つの領域ごとの平均値を先に求め、再度その4つの値を平均して求められる。1つの文字内で色が変わった場合は、上述の文字コードへの変換は行わない。

## 【 0 0 1 8 】

次に、文字除去部 2 1 4 では、文字認識部 2 1 2 で認識された文字画像データを元の L a b 成分の画像データから除去する（S 1 6）。元の L a b 成分の画像データは、上述の O C R 前処理（S 1 2）における 2 値化処理において別の場所に保存されているものである。文字除去処理において、具体的には、2 値化した画像データに対し、図 9 に示す 5 × 5 の最小値フィルタを用いるフィルタ処理を行うことで、文字画像データを膨張させる。膨張させた文字画像データに対応する L a b 成分の画像データ部分を白抜きする。

## 【 0 0 1 9 】

次に、白抜きした L a b 成分の画像データ部分を、図 1 0 に示す 7 × 7 のフィルタによってフィルタ処理を行うことによって周辺画素の値に応じて補完する（S 1 8）。図 1 0 に示す注目画素は、白抜きされた部分の画素であり、フィルタ内に示されている番号は周辺画素を参照する順番を示している。フィルタ内の順番に従って、注目画素の周辺を縦横斜めの 8 方向について、白でない画素があるか否かを調べていき、はじめから 3 つ目までの白でない画素の値の平均値で注目画素を補完する。このようにして、文字画像のない背景画像のみの画像データが作成される。さらに、画像データ量を減らすために画像データの必要部分のみを切り出して保存する。図 1 1 に示すように、背景画像のみの画像データを左上から順に走査し、所定の閾値以上の濃度を持つ領域を矩形のビットマップ画像データとして切り出す。

## 【 0 0 2 0 】

以上のようにして作成された背景画像のみのビットマップ画像データと、文字認識部 2 1 2 で認識された文字コードデータとを、図 1 2 に示すように配置し合成する（S 2 0）。合成方法は、出力するファイル形式に応じて様々であるが、例えば、R T F や P D F の形式で出力する場合、画像データと文字コードはそれぞれ別々のオブジェクトとして扱われるので、画像を下に配置し、その上に文字コードデータを上書きすることで、図 1 2 に示すように合成される。

## 【 0 0 2 1 】

なお、本実施形態ではスキャナ 4 0 0 で読み取って得られた画像データに対し

てコンピュータ 2 0 0 で処理を行っている。しかし、同様の処理をスキャナ 4 0 0 内で行ってもよいし、プリンタ 3 0 0 で行うようにしてもよい。また、スタン  
ドアロンのデジタル複写機のような形態にも適用できる。

## 【 0 0 2 2 】

### 【発明の効果】

背景画像のある文字を、文字を除去した背景ビットマップ画像と文字コードに  
分離して再合成するので、

( 1 ) コード化した文字と原稿の文字画像が 2 重に出力されることなく、原稿  
の背景を再現した出力が得られる。

( 2 ) 背景のある文字でも、文字がコード化されるので、読みやすい出力が得  
られ、文字サイズの変更ができ、また、検索に利用できる。

( 3 ) 背景のデータも再利用できる。たとえば、文字を変更しても、同じ背景  
の文書が作成できる。

色変化のある文字画像データは文字コードデータに変換しないことにより、文  
字以外のものを文字とする誤認識を減らせる。

### 【図面の簡単な説明】

【図 1】 原稿と従来 of 出力例の図

【図 2】 原稿と従来 of 他 of 出力例の図

【図 3】 システム構成を示す図

【図 4】 本発明 of 画像データ出力 of 1 例 of 図

【図 5】 画像処理全体 of フローチャート

【図 6】 文字 of 外接矩形 of 図

【図 7】 最大値フィルタ of 図

【図 8】 色検出を説明するための図

【図 9】 文字除去 of 処理を示す図

【図 1 0】 補完データを探す画素位置 of 順 of 図

【図 1 1】 画像データ切り出しを説明するための図

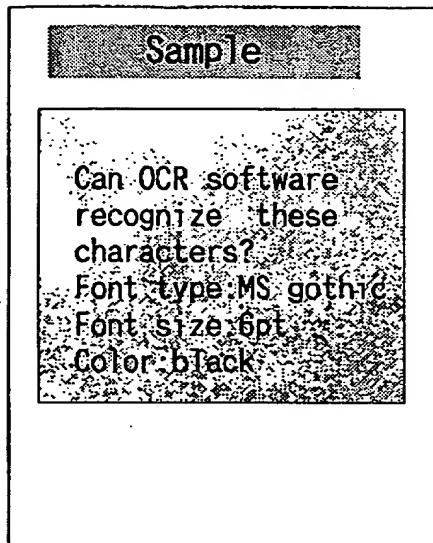
【図 1 2】 合成画像 of 図

### 【符号 of 説明】

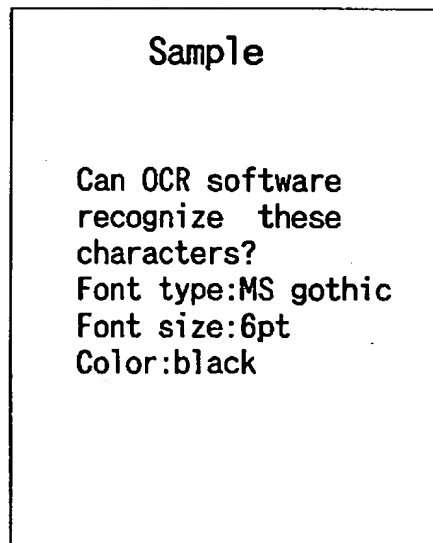
100 ネットワーク、 200 コンピュータ、 205 記憶装置、  
210 文書認識処理部、 300 プリンタ、 400 スキャナ。

【書類名】 図面

【図1】

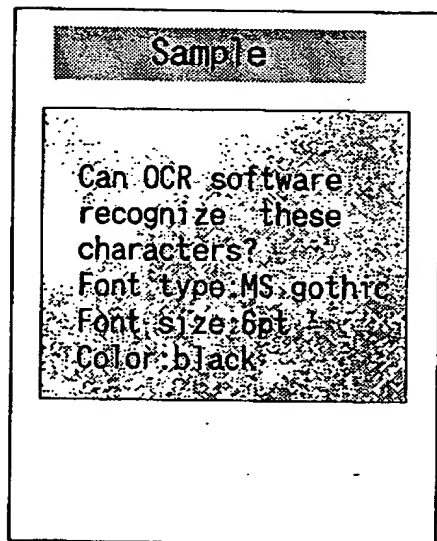


原稿

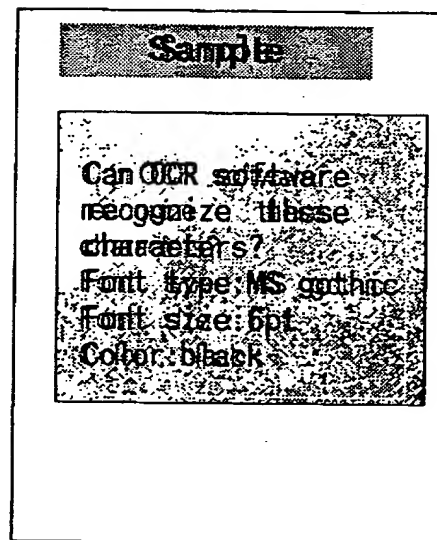


出力例

【図2】

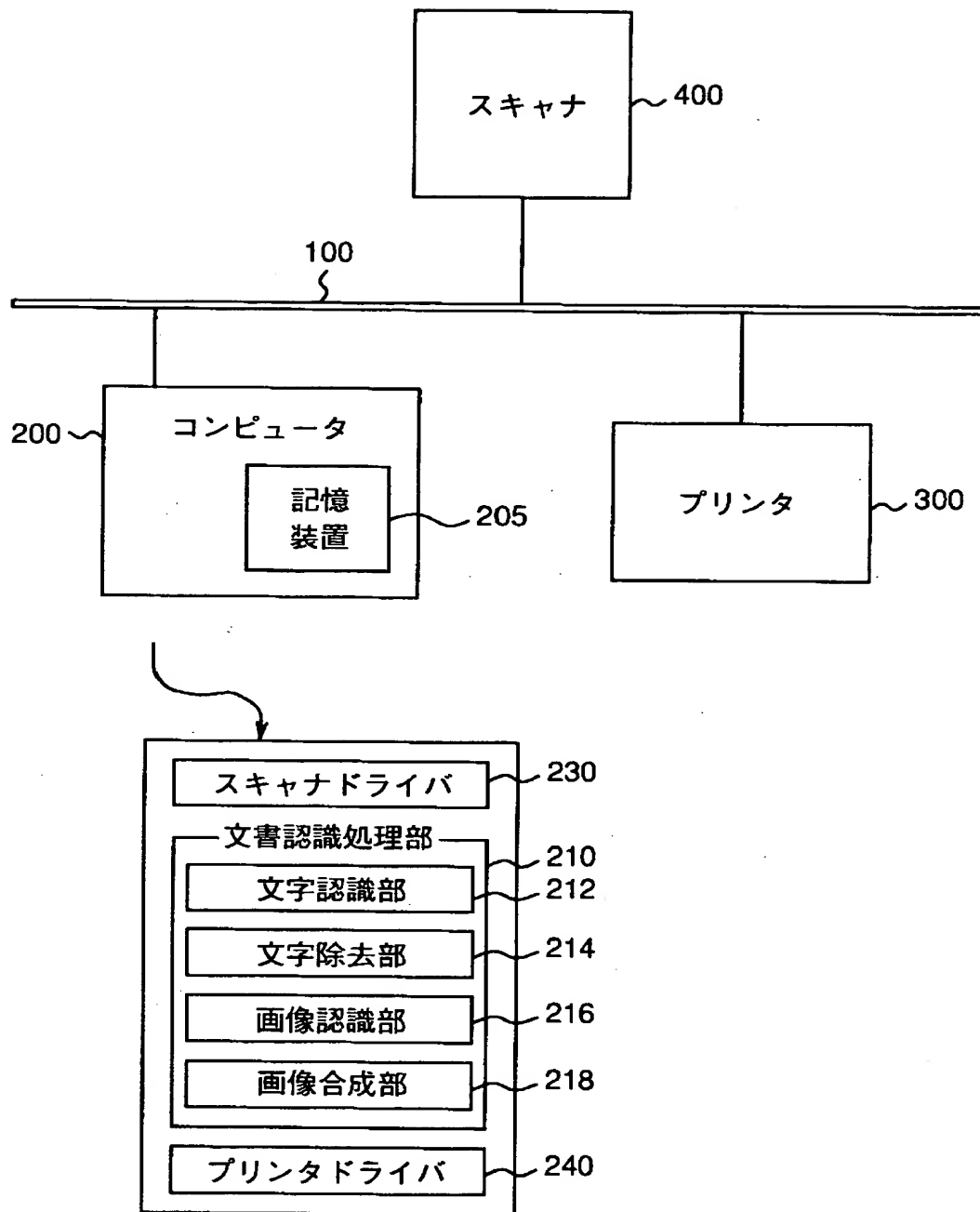


原稿



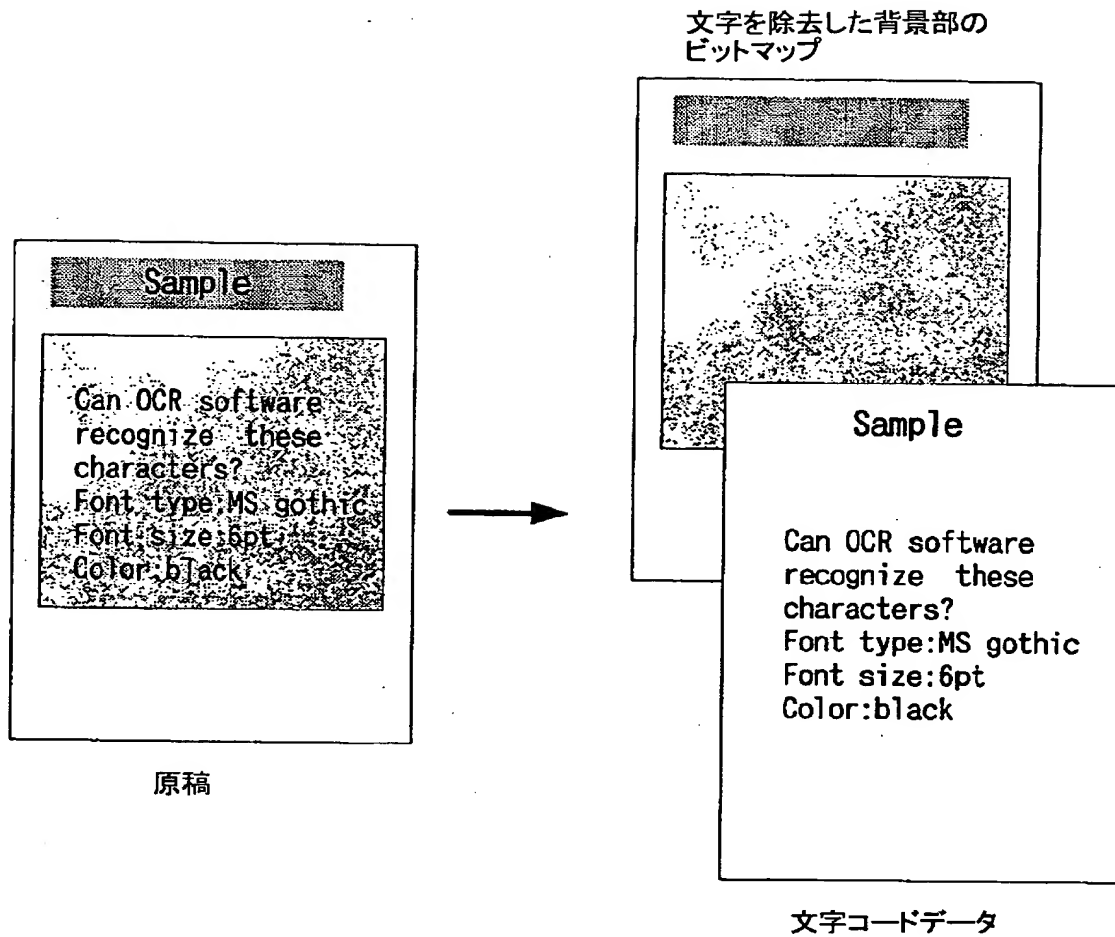
出力例

【図3】

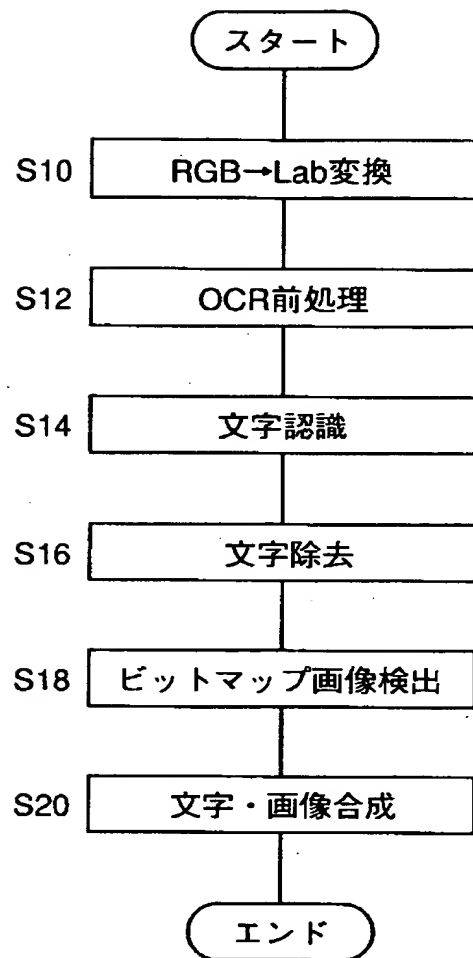




【図 4】

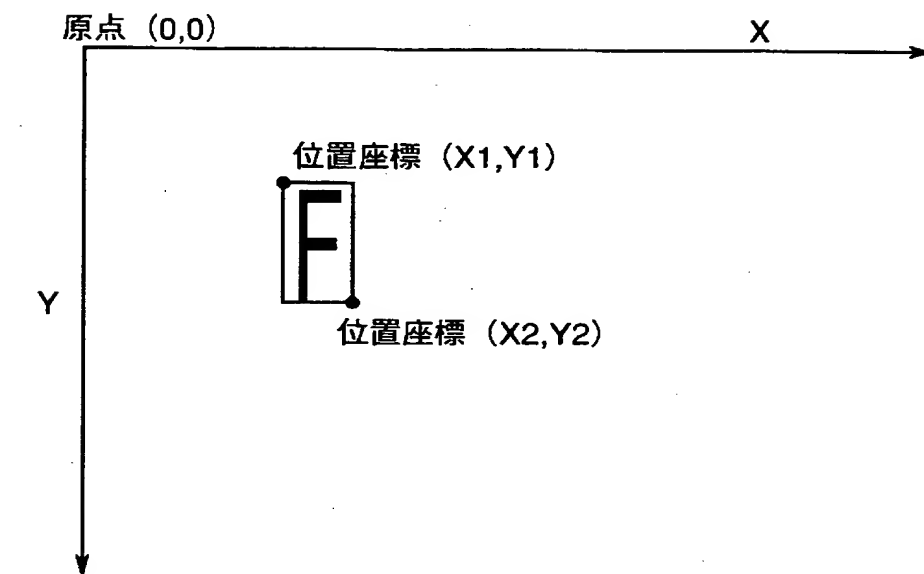


【図 5】



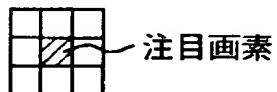
【図 6】

Font type

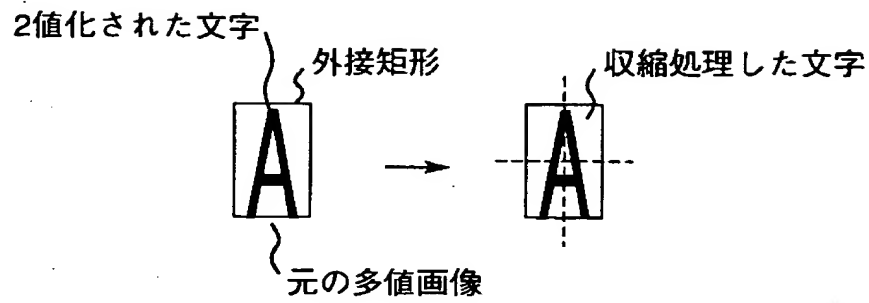


【図 7】

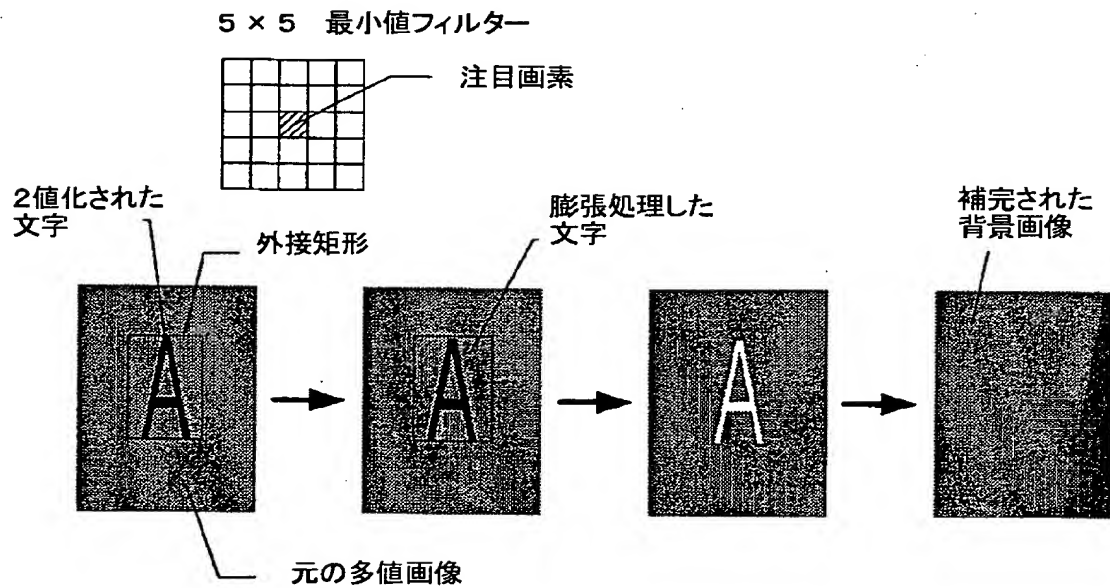
3×3 最大値フィルター




【図 8】



【図 9】



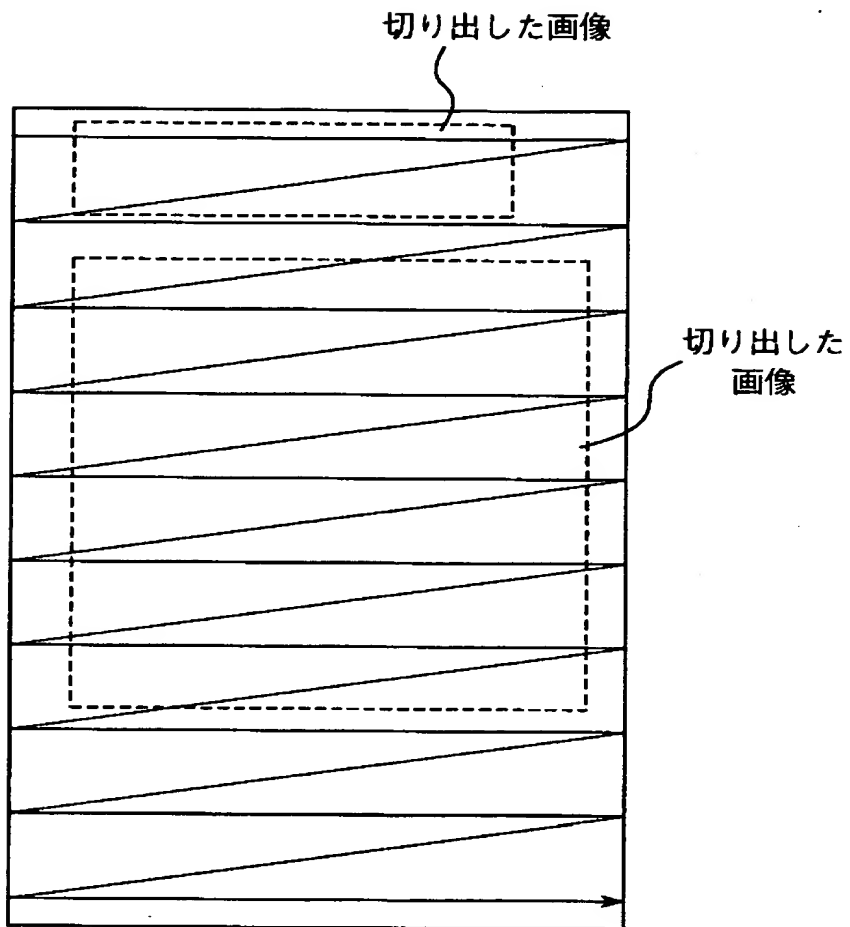
【図 1 0】

24			17			21
	16		9		13	
		8	1	5		
20	12	4		2	10	18
		7	3	6		
	15		11		14	
23			19			22

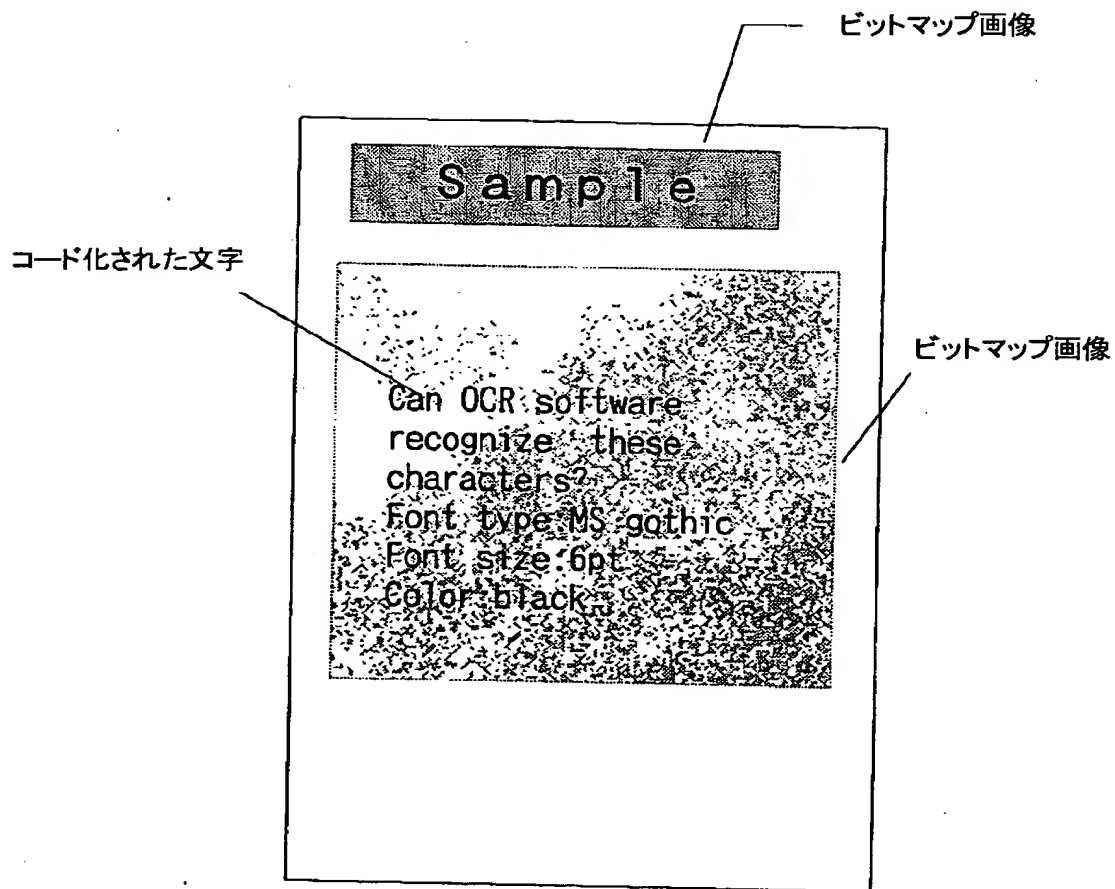


注目画素

【図 1 1】



【図12】



【書類名】 要約書

【要約】

【課題】 画像処理において背景のある文字をうまく再現する。

【解決手段】 文書処理において、文字を含む入力多値画像データから、2値画像データを作成し、この2値化データから文字コードと文字位置を認識する。次に、認識された文字の位置データを用いて多値画像から文字を除去する。そして、文字を除去した画像データとコード化された文字データとを重ねて配置して、合成画像を出力する。

【選択図】 図5



出 願 人 履 歴 情 報

識別番号 [000006079]

1. 変更年月日	1994年 7月20日
[変更理由]	名称変更
住 所	大阪府大阪市中央区安土町二丁目3番13号 大阪国際ビル
氏 名	ミノルタ株式会社